Proc. Ind. Assoc. Cult. Sci. **84A** (1), p1-p11 (2001)

IJP A
— an international journal

Proceedings of IACS

New opportunities in biological and chemical crystallography*

John R Helliwell
Department of Chemistry, University of Manchester,
M13 9PL, England
E-mail : hell@spec.ch.man.ac.uk

Abstract : K Banerjee (1933) offered a new way to approach the crystallographic phase problem which not only broke new ground beyond the 'trial and error' structure solution method of that time, but also heralded the extremely powerful direct methods of crystallography of the modern era from the 1970s onwards in chemical crystallography. Some 200,000 crystal structures are known today More complex crystal structures such as proteins required new experimental and theoretical methods to solve the phase problem. These are still evolving and new methods and results involving synchrotron radiation at softer X-ray wavelengths (2Å) are reported. In addition an overview is given of the new opportunities that are possible for biological and chemical crystallography especially via harnessing SR and neutron beams

Plan of the Lecture

First Endowment Lecture for Professor K. Banerjee delivered at the Indian Association for the Cultivation of Science, Calcutta, India on 19 September 2000.

1. Introduction and cross-disciplinary connections : from physics to biology and chemistry

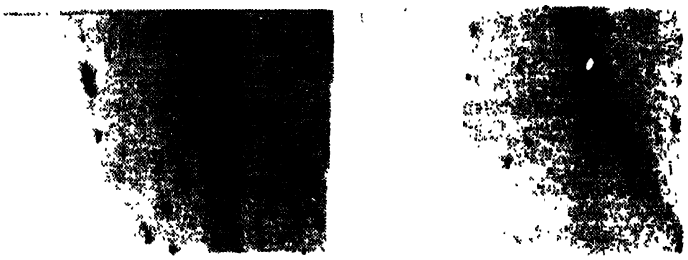
New opportunities in biological and chemical crystallography arising from physics based developments in synchrotron radiation and neutron beam production are at an exciting stage. The synergy between developments in one area of crystallography and another is of great interest. This article describes these synergistic developments. It is an honour to be awarded the Professor K. Banerjee Memorial Lecture Silver Medal as part of the birth centennial celebrations for this pioneer of crystallography. To gauge the honour, I need only cite his seminal paper on direct methods [1] *i.e.* pre-dating by nearly 20 years the next papers on direct methods and which led to the Nobel Prize in Chemistry in 1985 to H. Hauptmann [2] and J. Karle [3]. The compendium of Banerjee's publications [4] shows breadth too covering experimental techniques in crystallography and diffraction, structural chemistry and, as referred to above, the mathematical and physical basis of diffraction and the crystallographic phase problem. Another famous connection for Professor Banerjee in the history of crystallography is that he worked with Sir W.H. Bragg in the 1930s. W. H. Bragg, winner of the Nobel Prize in Physics in 1915 with his son W. L. Bragg, invented the four-circle diffractometer, a most famous example of physics and instrumentation furthering crystallographic science. This historical context, of synergies between scientific fields, shows that cross-disciplinary connections in crystallography were common then, as they are now. This article explores modern examples of such synergies. Figure 1 shows photos of Professor K. Banerjee.

2. Anomalous scattering and tuneable synchrotron radiation (SR); recent examples involving softer X-rays

The phase problem of crystallography arises because it is not possible to record the phase of a 'Bragg reflection' with respect to the incident beam. The desired experimental 'phase sensitive' detector at X-ray frequencies does not exist. Instead, some way has to be found from the measured intensities alone to obtain the phases which are then used in the Fourier series summation to obtain the electron density distribution in the crystal unit cell. Banerjee's insight [1] that an 'algebraic relationship' between the phases of several reflections might exist is the core idea of the 'direct methods'. Banerjee, in a review written in 1960 [5], described direct methods as either algebraic (*i.e.* like his approach) or statistical *i.e.* probabilistic (like the methods of Karle and Hauptmann [2, 3]). The two are very closely connected in principle. Both approaches suffer from being "too weak' to be applied to large structures; Banerjee certainly did not pursue his approach, as he said [5], for that reason. Crystal structures involving the placement and assignment of upto 200 atoms are now routine; some 200,000 chemical crystal structures have been solved today. Is there such a possibility for complex structures *i.e.* of proteins? Currently there are 12,000 protein crystal structures in the Protein Data Bank (PDB), of which ≈ 2000 are distinct structures (*i.e.* excluding genetic mutants or related crystal forms). The chemistry based phasing approach of multiple isomorphous replacement (MIR) of heavy atom substitution and multiple data set measurements from several crystals was developed by Kendrew and Perutz in Cambridge through the mid to late 1950s and was the mainstay of 'de novo' protein crystal structure determination in the 1960 – 1990 period. This situation has changed with the advent of polychromatic SR where a multiple wavelength anomalous dispersion approach, based on measurements from a single crystal of a protein containing a suitable anomalous scatterer, becomes practical. Such a formalism, based on a minimum of two wavelengths (λ_1 and λ_2) and three experimental measurements ($F\lambda_1^+$, $F\lambda_1^-$ and $F\lambda_2^+$ or $F\lambda_2^-$, where (+) refers to an *hkl* reflection and (–) refers to its inverse reflection *hkl* or their crystal symmetry equivalents) was described as early as 1956 by Okaya and Pepinsky [6]. An 'algebraic approach' by Karle [7] and Hendrickson [8] or a treatment analogous to MIR (see Helliwell [9, 10]) have made the 'one crystal' approach very effective. This is readily applicable due to two factors; firstly some 40% of all proteins are metalloproteins, where the metal anomalous X-ray dispersion can be harnessed or, secondly, there is the possibility of preparing selenomethonine substituted proteins, where the Se K edge is well placed at $\lambda = 0.98$Å [11]. These two factors along with the steady, concerted, development of instruments at SR facilities capable of MAD measurements, has opened up the prospect of very large numbers of protein structures being determined, even on a genome scale *i.e.* one beam line alone delivering hundreds of protein structures per year already [12, 13].

These developments have been facilitated *via* the low emittance of the electron storage ring allowing high brilliance X-ray emission. This yields a finely focussed beam. Freezing of a protein crystal preserves the crystal sample against irradiation, although not indefinitely, usually allowing a full set of MAD data from one crystal. But with too big a crystal the sample can break up on freezing and the diffraction pattern is then spoilt (Figure 2). Hence, a fine focus beam from a high brilliance SR source is important. Small crystal size is also emerging as a regular feature of high throughput 'structural genomics' protein crystallography pilot projects where 'microcrystals' are cited as a frequent problem. Third generation high brilliance SR X-ray sources such as ESRF in Grenoble, APS in Chicago and SPRING-8 in Japan have reduced the sample size that can be used to the 20 micron, or less, size range, when working at 1Å wavelength. Further optimisation *via* the wavelength is possible *i.e.* to allow even smaller crystals to be studied ; the diffraction energy, E_{hkl} , into a reflection can be described by the following expression [14]

$$E_{hkl} = \frac{e^4}{m^2c^4\omega} I_0\lambda^3 LPA \frac{V_x}{V_0^2} |\underline{F}(h)|^2 . \qquad (1)$$



(a) (b)

Figure 2. (a) Correct choice of crystal sample size is important for freezing. Diffraction pattern (a) involved too big a crystal (0.5 mm to 0.8 mm in size, minimum to maximum) and yielded high mosaicity and split spots. (b) shows sharp, single spots (not split) from a crystal of 0.2 mm size and from which a full data set could be recorded and processed to 1.2Å. Work by J. R. Helliwell, P. Faulder and R. Nuttall at IMCA, APS, Chicago in collaboration with J. H. Naismith and D. Moothoo (sample disaccharide bound concanavalin A) Methods to reach atomic resolution but at room temperature, which is after all, closer to physiological temperature, seek to improve reflection peak to background by combining the SR beam near parallel collination with low mosaicity crystals [61] Hence, the link between crystal growth, data collection and structural detail is made [62]. Hence, one can readily see how high brilliance Sr X-ray sources, and their gradual evolution these last 20 years [63] has had such a profound effect on the data collection capability in protein crystallography structural studies.

Since the Lorentz factor, L, is proportional to $1/\lambda$ the overall proportionality on the wavelength, λ is λ^2. Thus softer X-rays can favourably enhance the diffraction from yet smaller crystals (*i.e.* small crystal volume V_x). There is a small penalty of utilising a longer wavelength in having a less favourable monochromator transmission polarization correction but overall an increase from 1Å to 2.5Å wavelength offers an order of magnitude increase in E_{hkl} , and which can thus allow a smaller crystal volume, V_x by an equivalent factor [15].

Figure 1. (a) Professor K Banerjee, born September 15, 1900 ied April 30 1975.

Figure 1. (b) Professor K. Banerjee in 1958 (age 58) with, on his left, Pandit Nehru and, next left, Indira Gandhi and others at the time of his Directorship of the Indian Association for the Cultivation of Science in Calcutta.

Use of such a wavelength, to enhance the diffraction from small crystals (Figure 3) is not then at, even near to, the Se K edge (0.98 Å), and is well beyond the K absorption edges of the commonly occurring transition metals (Fe, 1.743Å; Zn, 1.283Å, Cu 1.380 Å; Mn 1.896 Å *etc*) and beyond the L edges of the common heavy atom derivatives (Hg, L_{III} 1.000 Å; Au, L_{III} 1·04 Å; Pt L_{III} 1.073 Å *etc*). The crystallographic phase problem is likely to need new ideas in the 'structural genomics', small crystal, era to supplement the seleno MBD approach. Two new approaches are possible and offer exciting prospects. Firstly, the incorporation of xenon into a protein crystal under high pressure is possible [16]. Apparatus for freezing xenon in the protein crystal has been developed [17]. The 'xenon' diffraction data can then be recorded *e.g.* by MAD or, by combining with a native protein crystal diffraction data set, a hybrid phasing approach 'SIROAS' (Single Isomorphous Replacement with Optimised Anomalous Scattering) involving one suitably chosen wavelength is feasible also. Optimisation of the xenon anomalous signal using softer X-rays *via* the xenon L edge (L_1 = 2.27 Å) is possible [13]. This method has now been applied to the lobster apocrustacyanin Al protein using diffraction data recorded on the Daresbury SRS station 7.2 at 2 Å wavelength, the first protein crystallography station on the first dedicated SR source [18], now used in a novel way (Figure 4). Figure 4a shows the xenon anomalous difference Patterson and has peaks that are comparable, superior even than the xenon isomorphous difference Patterson (Figure 4b) by virtue of the 2 Å X-ray wavelength. A second benefit of softer X-rays is that the sulphur anomalous dispersion signal is enhanced, over that at 1.5 Å, by a factor of 2 even though the sulphur K edge optimisation at 5 Å wavelength is remote. Nevertheless, the prospect exists of determining the sulphur atom sub-structure from the disulphide atoms (at cysteine to cysteine covalent bridges between the polypeptide chain), and possible eg even from the single sulphur atoms in methionine amino acid residues. Such a substructure of the whole protein structure could be combined with atomic resolution protein crystal diffraction data and follow on from the successes of direct methods with metalloproteins [19] or solvent flattening based approaches [20]. Determination of the sub-structure of the anomalous scatterers is therefore a pivotal step. Larger and larger constellations of such atoms in the sub-structure necessitate automatic methods. A seminal paper was that of Mukherjee *et al* [21] who used anomalous differences and the direct methods program MULTAN.

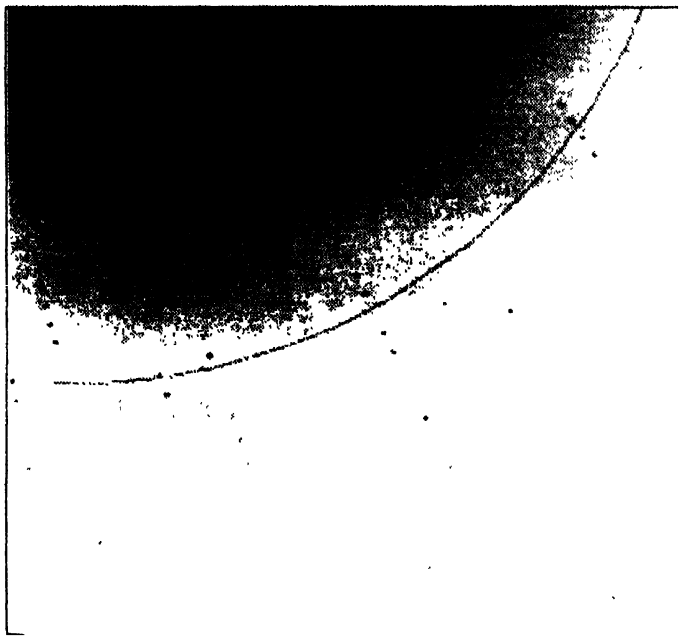


Figure 3. Softer X-rays have great potential in protein crystallography, especially 'structural genomics' projects. Very small crystal size is a common occurrence in current pilot projects. High quality diffraction data can be recorded using *e.g.* 2Å wavelength, evidenced here for apocrustacyanin A1 needle crystals ($\simeq$ 0.1 mm thick). From Ref [54].

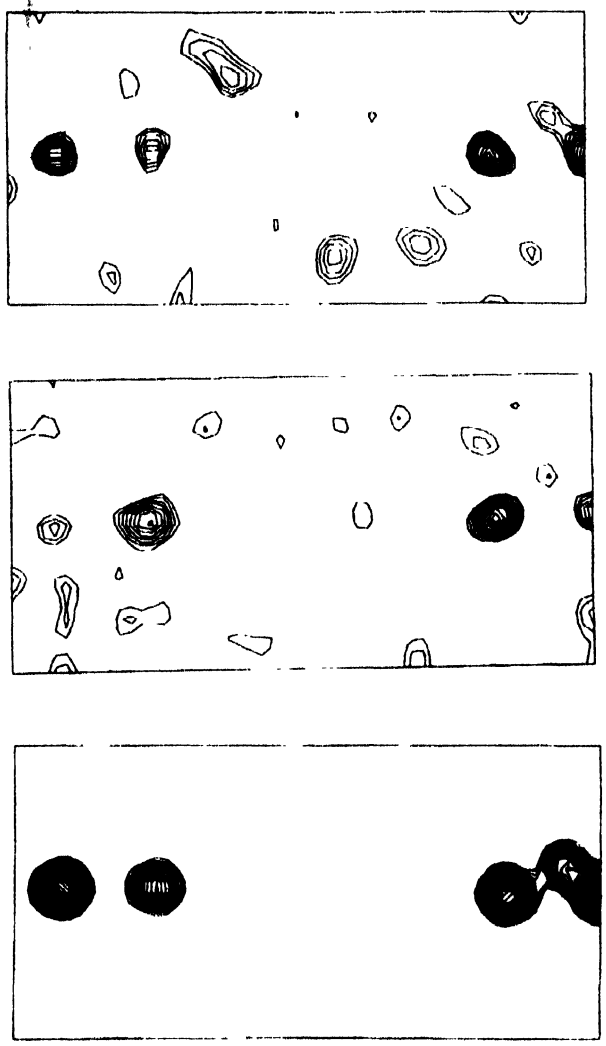


Figure 4. Softer X-rays also yield optimised anomalous scattering from the xenon L edge, and where xenon incorporated into a protein crystal under high pressure is now a commonly obtainable 'heavy atom' derivative. (a) Anomalous difference Patterson Harker section (b) Isomorphous difference Patterson Harker section for appocrustacyanin A1 as example Data recorded on station 7.2 at the Daresbury SRS, λ = 2Å [55]. The high quality of (a) is evident and quite comparable, superior even, with respect to the isomorphous case (b), (c) provides the calculated Harker section peaks from the xenon atom positions involved.

3. Laue diffraction and polychromatic SR also leading to synergy with neutrons

Unlike conventional X-ray sources which emit emission lines on a weak continuum the synchrotron X-radiation universal curve is a strong continuum of radiation. Hence extremely strong monochromatic beams can be extracted, for which the wavelength can be tuned (as referred to in the previous section). In this section, however, consideration is given to how to utilise all the wavelengths simultaneously. This is the well known, original, Laue method of crystallography of 1912. In this way ultra-short X-ray exposure times can be realised, as short as subnanosecond time resolution [22, 23]. An example of a time-resolved Laue enzyme structural study is described in Section 4.3.

An analogous situation to SR exists with neutron beams from a reactor source where a polychromatic spectrum of neutrons is emitted. The neutron Laue method allows the sample size barrier to be transcended, rather than achieving ultra-short exposure which is the purpose of the SR Laue method. With neutrons the size of a single protein crystal has been severely rate limiting. The harnessing of a broader band of wavelengths at neutron sources is an important development [24]. At reactor sources the wavelength bandpass that can be used is ≈ 33%. A vital next step for proteins beckons using the time-of-flight Laue approach, feasible at a proton synchrotron source, whereby both a broader bandpass can be used a better signal to noise realised, used so far in chemical crystallography [25]. Enhanced proton synchrotron power is coming, suitable for time-of-flight Laue protein crystallography with neutrons, *via* the proposed ISIS 2nd target station [26] and, ultimately the proposed European Spallation Source (5 MW *vs* 160 kW for ISIS) [27]. The importance of these developments for the future lies in the definition of the hydrogen atoms. Whilst SR X-rays have opened up the chance to visualise hydrogens for relatively well fixed atoms more mobile H atoms have a better chance to be defined when exchanged with deuterium and studied with neutrons. Hence, those catalytic hydrogens on more mobile pieces of polypeptide can be determined. Moreover, bound water molecules in ligand binding sites can be studied in more detail by neutron protein crystallography [28] and thus the chance to further the understanding of the thermodynamics of ligand binding to protein receptor sites *via* structure might be advanced [29, 30]. Structure based drug design is a pivotal method in finding new lead compounds in the pharmaceutical industry, and the science on which it is based is a vital topic in molecular biophysics and biophysical chemistry.

4. Synergy between protein and chemical crystallography

Three specific areas interfacing between large and smaller molecule crystallography will be highlighted *via* case studies.

4.1. Protein crystal structure analysis of concanavalin A with small molecule accuracy

Small molecule accuracy means that the protein crystal diffraction pattern extends to a resolution limit where the measurable number of unique reflections exceeds the number of atomic parameters (with full anisotropic ellipsoid refinement) by around an order of magnitude. Hence for a protein with 1800 atoms, some 180,000 X-ray reflections are needed to satisfy that criterion. There is a significant difference however between small and large molecule analysis because the protein has some mobile loops and side chains, which do not diffract well. These atoms require the benefit of standard model geometry restraints such as planar peptide bond and standard bond distances and angles. With this caveat in mind, small molecule accuracy protein refinement is possible today with the advent of cryo-protection of the protein crystal against radiation damage, strongly focussed SR beams and sensitive, automatic, area detectors. The crystal structure analysis of concanavalin A, initially at 0.94Å (with 117,000 unique X-ray data ; 20,000 model restraints and 18,500 parameters to be refined) [31] and extended to 0.92Å (with 20,000 more X-ray reflections) reached an X-ray data to parameter ratio of around 8:1 [32, 33]. Full matrix inversion using SHELXL [34] allowed standard uncertainties (s.u.) on bond distances to be determined and for well ordered parts of the protein (some 60%) s.u.'s of 0.005 to 0.01 Å were obtained. Thus, the precise metal (Mn and Ca) bioinorganic coordination was determined and also the protonation states of carboxylic acid side chains revealed from the X-ray analysis (typically the province only of neutron protein crystallography previously). Concanavalin A is a tetrameric protein which binds saccharide (Figure 5) and is involved in cell to cell cross-linking and is thought also to serve as an anti-fungal protection agent in jack beans.

Figure 5. Structural studies on concanavalin A from jack bean reveal a tetramer of protein subunits shown here in 'ribbon format', with two metal ions as small spheres and nearby glucose molecule, per subunit From Ref [33], based on Ref [29].

4.2. Multiple wavelength anomalous dispersion (MAD) : Mn, Ca concanavalin A and CoZn PO :

The MAD method is proving to be very effective as a means of

protein crystal structure determination especially involving selenomethionine expressed variants of a protein [8, 11–13]. Since a large percentage (estimated as high as 40%) of proteins are metalloproteins then the MAD method's potentiality to be applied widely to access the intrinsically present metal atom or atoms is possible. For metalloproteins quite a range of X-ray wavelengths are of interest eg from Mo K edge at 0.62 Å to, say, calcium K edge at 3 Å. Working with longer wavelengths does mean that X-ray absorption becomes a concern and smaller crystals need to be selected. Otherwise the principles of the method are similar. In the analysis two different procedures exist. One, the algebraic MAD method [7] relies on measurement redundancy and least squares (MADLSQ [35]). The other, second method, is a variant of multiple isomorphous replacement, with one wavelength serving as a native data set and the other remaining wavelengths (1, 2 or more) as derivatives, two wavelengths being the minimum (see Section 2). The map quality from adding more wavelengths has been evaluated with respect to a brominated oligonucleotide [36]. In such a case of a rather concentrated anomalous scatterer (2 bromines in 400 atoms) two wavelengths proved very satisfactory. The evaluation of map quality for the Mn, Ca metalloprotein concanavalin A has been undertaken by us using Mn K edge MAD data recorded at ELETTRA in Trieste [32]. With 1 Mn in 1800 atoms this was a more challenging case. The anomalous difference and dispersive difference Patterson maps both clearly revealed the Mn atom position and from phasing on that alone, the Ca atom position was found. The Mn MAD (including a small but significant calcium anomalous signal) with density modification phasing quality had a mean figure merit of 0.58 for three wavelengths (mean phase error of 55°) and the best two wavelength derived phase set had a mean figure of merit of 0.43 (mean phase error of 65°). Hence the need for three wavelengths was found to be necessary for such a case but the best two wavelength case was promising. Moreover, it was concluded, since frozen crystals of concanavalin A on the CHESS, Cornell multipole wiggler had yielded diffraction data to 0.92 Å resolution then a combined future structure solution approach involving MAD to around 2 Å resolution and phase extension to 0.92 Å resolution suggests itself for the future. Moreover, ultra intense and rapidly tuneable capability is now being combined in the one beamline, eg at ESRF [37], this will be feasible in one experimental data collection run. Also electron density maps will be calculable 'on the fly' as each data set at each wavelength and at the highest resolution are collected. High performance computing on the beamline is a necessity for that. This case study allowed the settling of issues re numbers of wavelengths and efficiency of beamtime usage posed over 20 years ago [9]. In terms of strong user centres worldwide for MAD the NSLS X4C beam line [38], CHESS, and ESRF BM14 [12] and LURE [39] are examples. On BM14 at ESRF for example over a hundred MAD protein structures are solved typically per year.

The development and application of MAD techniques in inorganic (smaller molecule) chemical crystallography has been reviewed recently [40]. From the Manchester Laboratory, again using the ELETTRA 'XRD' station, a five wavelength MAD analysis has been conducted on the zincophospate, CoZnPO (CZP) [41]. This analysis allowed the site of incorporation of cobalt to be settled, *i.e.* at which of the two possible (or both) zinc sites. The most sensitive wavelength pair to the cobalt occupancy was the 1.45 Å reference wavelength and the f'' minimum wavelength (1.608 Å) between which difference structure factor amplitudes were calculated. A complication versus a protein MAD typical case is that there were sufficiently strong variations in the structure factor amplitudes, due to the zinc atoms' contributions, to complicate the scaling of the data sets between the wavelengths. However, since an atomic model for the crystal structure was available (only the cobalt substitution effect was to be settled) it was feasible to calculate which structure factor amplitudes had a small zinc atoms contribution (indeed there were not so many of these). Hence this small subset of reflections could be used to put the data sets on a common scale. The analysis concluded by estimating that such a MAD analysis would be sensitive down to 12 - 15% cobalt substitution levels. Overall, these techniques are applicable to settling *details* of a small molecule crystal structure rather than structure solution *per se* (which is usually straightforward by conventional Patterson or direct methods). Such special cases are not uncommon however for instance in metal substituted alumino phosphates as well as compounds of several heavy atoms, of close atomic number, and where wavelength tuning leads to unambiguous atom identification. Yet more sophisticated analyses combining MAD with change of temperature has been reported and for which metal ion migration can be tracked in the crystal [42].

4.3. Time-resolved and temperature-resolved crystallography : hydroxymethylbilane synthase (HMBS) enzyme and a liquid crystal study respectively .

The study of the evolution of structural intermediates with time of reactive protein molecules such as enzymes like HMBS and the perturbation of smaller molecules with temperature as a variable is a modern growth trend in crystal structural analyses. These developments have been made feasible due to the hugely expanded capabilities for fast, repeated, data set collections. Thus a structure to function relationship can be explored directly by experiment. A particular impetus in time-resolved protein crystallography has arisen from harnessing the Laue method with focussed polychromatic SR beams where even sub-nanosecond exposures are possible [22]. Thus, the range of time-resolutions reaches well beyond freeze trapping of structural intermediates. Time-resolved techniques have been applied in biological crystallography eg see the ten case studies over the last decade highlighted in Ref. [23]. In chemical crystallography, where the scattering strength [eq. (1)] is higher, due to the smaller unit cell volume, ultra-fast monochromatic techniques have been preferred so far over the polychromatic technique [42]. In a crystal, both reversible and irreversible structural change can be stimulated. Reversible cases are amenable to stroboscopic (cyclic) data collection.

HMBS catalyses the polymerisation of four molecules of porphobilinogen to form hydroxymethylbilane. The evolution of the reaction in crystals of a Lys 59 Gln mutant was studied by repeated data collection *via* Laue diffraction snapshots. A progression of Laue exposures with pre-set, lengthening, time gaps after substrate solution passed over a crystal held in a flow cell were recorded at ESRF on the 'Laue beamline'. The experimental Laue difference maps revealed an elongated difference electron density peak, most prominent after 2h, commencing after ≈ 8 minutes, and by which time the initially colourless enzyme crystal had become red/pink [43]. This peak commenced at the position of C2 of the oxidised co-factor of the enzyme (the putative binding site for substrate) and directly above the critical carboxyl side chain of Asp 84 involved in the first ring coupling reaction step. The electron density then extended past amino acid residues that are known from protein engineering to affect later stages of the catalysis, and into open solvent. There is a missing loop of residues (49 – 57) in all current HMBS structures undertaken at ambient temperature. The Laue derived density resides between where that loop is likely to be and the position of the C2 ring in the reduced active cofactor determined by MAD [44]. Thus,the time-evolution of this reaction in the HMBS (Lys 59 Gln mutant) crystal in structural and functional terms has been established over a period of seconds to hours [45]. Structural puzzles remain in the study of this fascinating and important enzyme. For example, repeated ring coupling is required in the tetrapyrrole formation but there are not four conserved carboxyl side chains (akin to Asp 84). Domain: domain reorganisation has been proposed [46]. Further studies involving time-resolved small angle scattering in solution of the enzyme have been suggested [45] to follow the enzyme radius of gyration with time as substrate is mixed with the enzyme. In the crystal of course, any domain to domain rearrangements would be inhibited.

A temperature resolved chemical crystallography study utilising SR sources has been undertaken on a thermotropic liquid crystal 1, 4, 8, 11, 15, 18, 22, 25-octahexylphthalocyaninato nickel [47]. This study has been in two parts [47, 48] and involves a sequence of crystal structures at 100 K, 293 K (*i.e.* room temperature), 323 K, 328 K and 353 K, there being a sudden structural change between 323 K and 328K. A complication of the liquid nitrogen temperature study at 100 K was that there was a crystal space group transition between 195 K and 205 K. As expected the most precise structure "snapshot' is the lowest temperature one. The evolution of these structures with temperature increase shows an increased thermal motion of the hexyl groups. These motions presumably get so large as the actual liquid crystal transition temperature of 428 K is approached that free relative rotation of the molecules can take place, to give the discotic columnar mesophase *i.e.* where there is random rotational orientation of the molecules about the symmetry axis of the phthalocyanine ring. Conversely at the lower temperatures, the phthalocyanine cores are locked together by the ordering of these hexyl groups. Hence, one sees directly here a neat explanation of the structural basis of this important phase transition from the crystal to liquid crystal molecular phases. The data collection for these studies involved the use of the 24 pole wiggler beamline station F2 at the CHESS synchrotron (for the most precise low temperature structure at 100 K, as well as at 293 K). The elevated temperature studies were undertaken on the SRS, Daresbury wiggler beamline station 9.8. The crystal space group transition was determined on SRS Daresbury Station 7.2 with a CCD time slicing detector [47].

5. Future directions

In macromolecular crystallography there are important current trends involving the elucidation of large multi-macromolecular complexes, structural genomics and a growth in neutron protein crystallography. An example of a large complex involves the structures of the ribosomal components which are being reported after many years of effort overcoming technical difficulties [49, 50, 51]. In another area, the architecture of symmetrical viruses has of course been the subject of intense study from the late 1970s onwards [52]. A most interesting multi-macromolecular complex under investigation is the 16mer protein subunit complex involved in the bathochromic spectral shift of bound carotenoid in the crustacyanin from lobster and Vellela vellela. This has apparently evolved as a camouflage mechanism of marine crustacea against predators. The elucidation of the crystal structures is underway [53–57].

The rate of determinations of protein crystal structures has improved to such an extent that high throughput approaches are being considered at the genome numbers scale. Amongst the first such programmes, in effect a pilot project, is that of Sung Ho Kim's in Berkeley. The feasibility of this has rested on MAD beamline availability to use the intrinsic metal atoms in metalloproteins and/or the selenium introduced *via* seleno-methionine bacterial growth. A bottleneck reported in these pilot programmes, which are based principally in the USA at present, is the occurrence of microcrystals for many of the expressed proteins. This is understandable since the time to optimise conditions is reduced compared with 'standard' protein crystallographic projects. An additional way, besides the high brilliance of SR, to overcome small crystal, weak, diffraction is to use softer X-rays and thus enhance the scattering efficiency [eq. (1)]. This is then beyond the Se K edge and the transition metal absorption edges as well as the L edges of the standard isomorphous derivatives used in MIR phase determination. Instead, the xenon derivatisation approach allows for optimisation of the anomalous scattering *via* the softer X-rays too (as discussed in Section 2). Overall, high throughput protein crystallography should lead to a large increase in the database of precise protein structures in the Protein Data Bank. This will help provide a platform of protein structures as data for many areas of molecular biophysics, biophysical chemistry, folding and structure prediction, and not least in opening new avenues in lead compound discovery towards new pharmaceuticals. The large scale data mining of such information is akin to that exists

already in chemical crystallography with its current database of over 200,000 chemical crystal structures. This is then another example of the intersection of chemical crystallography and protein crystallography.

The potential of neutron protein crystallography is strong with a number of initiatives underway. At the neutron reactor source in Grenoble, the most powerful in the world, there is a coordinated millennium instrument refurbishment going on. This includes upgrades to the relevant instruments namely D19 (a monochromatic neutron diffractometer with enlarged area detector coverage (thus improving data collection efficiency) and the Laue Diffractometer (LADI) which will have a new neutron image plate reader and a higher flux location yielding gains of 3 to 5 times in sensitivity + flux (thus reducing data collection time or allowing smaller crystal samples to be used or larger molecular weight proteins to be investigated or combinations thereof). Thus, the determination of hydrogen to deuterium exchange details can be applied to a wider variety of proteins than ever before. The determination of the hydrogens themselves is now feasible with ultra-high/atomic resolution protein SR X-ray crystallography where these hydrogens are well defined. Mobility of hydrogens can kill their diffraction signal however but, since neutron protein crystallography determination of deuteriums at around 2Å or better resolution matches that at 1.0 Å by SR X-rays, then more mobile hydrogens are determinable by the neutron approach. The bound solvent is a whole category of deuterium atoms which are more efficiently sought by neutron techniques [28]. Full deuteration through appropriate microbiological expression of proteins for bacteria grown on deuterated media is possible ; it was recently shown [58] how this worked for deuterated myoglobin, where a 1.7 Å neutron study was more effective than a 1.5 Å X-ray study in finding even the relatively static hydrogens (as deuteriums). In the short to medium term, there is the proposed UK ISIS 2nd Target Station [26] which is set to produce longer wavelength neutrons which scatter more strongly (*i.e.* as per equation 1) and there is included in that development the plan for a large molecule crystallography time-of-flight instrument (LMX). In USA, there is their Spallation Neutron Source (SNS) under construction at Oak Ridge which at 2 MW is much more powerful than ISIS at 160 kW and, pro rata, delivers more flux. A neutron protein crystallography time-of-flight station on the SNS is being discussed. In Europe, there is the proposed 5 MW European Spallation Neutron Source (ESS) envisaged for construction beginning in 2010 [27]. The proposed ESS would reach the same average flux as the Grenoble reactor neutron beam flux but with the addition of the benefit of the time-of-flight neutron approach. Hence, the detailed dissection of those enzyme mechanisms involving hydrogen atoms, of which there are many, is set for an expansion in examples.

Overall, the experimental capabilities for protein crystallography analyses have changed dramatically in scope through a combination of technology and advances in methods [59], from both the physical sciences and the biological sciences. Trends seen in chemical crystallography of efficient structure determination and the vast expansion of structural databases is set to occur for proteins. The application of developments in protein crystallography to chemical crystallography such as small crystals, SR MAD and perturbation crystallography is an exciting cross fertilisation [60]. Crystallographic science is a highly cross-disciplinary science. The years of Professor Banerjee, as witnessed by his compendium of publications in crystallography [4], show a predominant interaction between physics and chemistry. In the last decades, that has expanded strongly also to include biology.

Acknowledgment

I am very grateful to my past and present colleagues and students in the University of Manchester Structural Chemistry Research Laboratory, which is devoted to biological and chemical crystallography and includes extensive crystallographic methods development. I especially thank Dr Madeleine Helliwell, Dr James Raftery and Dr Jarjis Habash for their long term collaboration, and indeed all my co-authors of scientific publications over the years for their collaboration The synchrotron radiation facilities at Daresbury, ESRF (Grenoble), Cornell, NSLS and Elettra (Trieste) as well as the neutron beam facility in Grenoble (Institut Laue Langevin) have all provided splendid X-ray and neutron beams in support of this research programme in its various aspects over the years and covered in this article, and to whom JRH is also very grateful. Funding support from EU Network Crystallogenesis Project #BI04-CT98-0086 is acknowledged for the work referred to in Figures 2, 3 and 4. The research grant support of BBSRC (for computing facilities), The Wellcome Trust and The Leverhulme Trust and of EPSRC, BBSRC, British Council and the UK/Israel Fund for PhD studentship support are all also gratefully acknowledged by JRH. The Royal Society funded JRH with a conference travel award to Calcutta for the Professor K Banerjee Birth Centennial Conference celebration held in September 2000.

References

[1] K Banerjee *Proc Roy Soc.* **141** 188 (1933)

[2] H A Hauptmann (1985) *Nobel Lectures in Chemistry* (Singapore . World Scientific) (1992)

[3] J Karle (1985) *Nobel Lectures in Chemistry* Published 1992 (Singapore : World Scientific) (1992)

[4] K Banerjee *Collected Works*, (Indian Association for the Cultivation of Science, Calcutta 700 032) (2000)

[5] K Banerjee *Physics Teacher* **2** 5 (1960)

[6] Y Okaya and R Pepinsky *Phys Rev* **103** 1645 (1956)

[7] J Karle *Physics Today* **42** 6 22 (1989) (and references therein)

[8] W Hendrickson *Science* **254** 51 (1991)

[9] J R Helliwell *Daresbury Study Weekend Proceedings* DL/SCI R13, p1 (1979)

[10] J R Helliwell Macromolecular *Crystallography with Synchrotron Radiation* (Cambridge : Cambridge University Press) (1992)

[11] W Hendrickson, J R Horton and D M Lemaster *EMBO J* **9** (5) 1665 (1990)

[12] A Cassetta, A M Deacon, S E Ealick, J R Helliwell and A W Thompson *J. Synchrotron Rad.* **6** 822 (1999)

[13] N E Chayen and J R Helliwell *Physics World* **5** 43 (1998)

[14] M M Woolfson *X-ray Crystallography* 2nd edn (Cambridge . Cambridge University Press) (1970)

[15] J R Helliwell in *Daresbury Study Weekend CCP4 Proceedings* (Daresbury Laboratory) p 80 (1993)

[16] B P Schoenborn, H C Watson and J C Kendrew *Nature* **207** 28 (1965)

[17] O Sauer, A Schmidt and C Kratky *J. Appl. Cryst* **30** 476 (1997)

[18] J R Helliwell, T J. Greenhough, P Carr, S A Rule, P R Moore, A W Thompson and J S Worgan *J Phys* **E 15** 1363 (1982)

[19] M Mukherjee *Acta Cryst.* **D55** 820 (1999)

[20] Z Dauter, M Dauter, E De La Fortelle, G Bricogne and G Sheldrick *J Mol Biol* **289** 83 (2000))

[21] A Mukherjee, J R Helliwell and P Main *Acta Cryst* **A45** 715 (1989)

[22] D Bourgeois, T Ursby, M Wulff, C Pradevand, A LeGrand, W Schildkamp, S Laboure, V Srajer, T-Y Teng,, M Roth and K Moffat *J Synchrotron Rad* **3** 65 (1996)

[23] Z Ren, D Bourgeois, J R Helliwell, K Moffat, V Srajer and B L Stoddard *J Synchrotron Rad.* **6** 891 (1999)

[24] J R Helliwell and C Wilkinson HERCULES Book Chap XII "*X-ray and Neutron Laue Diffraction*" **Vol III** p 212 (Les Editions de Physique, Springer Verlag) (1997)

[25] C C Wilson and R I Smith in *Time-Resolved Diffraction* (eds.) J R Helliwell and P M Rentzepis (Oxford · Oxford University Press) p 401 (1997)

[26] Rutherford Appleton Laboratory Report, RAL-TR-2000-032 "*A Second Target Station at ISIS*"

[27] European Science Foundation Report ESF **Vol I** and **II** "*The European Spallation Source*" (1997)

[28] J Habash, J Raftery, R Nuttall, H J Price, C Wilkinson A J Kalb (Gilboa) and J R Helliwell *Acta Cryst.* **D 56** 541 (2000)

[29] G M Bradbrook, T Gleichmann, S J Harrop, J Habash, J Raftery, A J Kalb (Gilboa), J Yariv, I H Hillier and J R Helliwell *J Chem Soc. Farad. Trans.* **94** 1603 (1998)

[30] G M Bradbrook, J Forshaw and S Perez *Eur J. Biochem* **267** 4545 (2000)

[31] A Deacon, T Gleichmann, A J Kalb (Gilboa), H J Price, J Raftery, G Bradbrook, J Yariv and J R Helliwell *J. Chem Soc. Farad. Trans.* **24** 4305 (1997)

[32] N S Hunter, T J Gleichmann, M Helliwell, L Barba, E Busetto, A Lausi, H J Price and J R Helliwell *Croatica Chemica Acta* **72** 673 (1999)

[33] A J Kalb (Gilboa), J Habash, N S Hunter, H J Price, J Raftery and J R Helliwell in *Metal Ions in Biological Systems* **Vol 37** "*Margarese*" (ed) A & H Sigel Chap. 9 p 279 (2000)

[34] G Sheldrick *SHELXL 96* (University of Gottingen) (1996)

[35] W Hendrickson *Trans. Am. Cryst. Assoc.* **21** 11 (1985)

[36] M R Peterson, S J Harrop, S McSweeney, G A Leonard, A W Thompson, W Hunter and J R Helliwell *J. Synchrotron Rad.* **3** 24 (1996)

[37] S McSweeney and A W Thompson (*Personal Communication*)

[38] W Hendrickson *J. Synchrotron Rad.* **6** 845 (1999)

[39] R Fourme, W Shepard, M Schilz, T Prangé, M Ramin, R Kahn, E. de la Fortelle and G Bricogne *J. Synchrotron Rad.* **6** 834 (1999)

[40] M Helliwell *J. Synchrotron Rad.* **7** 139 (2000)

[41] M Helliwell, J R Helliwell, V Kaucic, N Zobukovec Logar, L Barba, E Busetto and A Lausi *Acta. Cryst.* **B55** 327 (1999)

[42] P Coppens in "*Twenty Years of Synchrotron Radiation Crystallography*" (eds) S Ealick and J R Helliwell (American Crystallographic Association Transactions) No 34 (1999)

[43] J R Helliwell, Y P Nieh, A Cassetta, J Raftery, A Hädener, A C Niemann, A R Battersby, P D Carr, M Wulff, T Ursby, J P Moy and A W Thompson in Time-Resolved Diffraction (Eds) J R Helliwell and P M Rentzepis (Oxford . Oxford University Press) p 187 (1997)

[44] A Hädener, P K Matzinger, A R Battersby, S McSweeney, A W Thompson, A P Hammersely, S J Harrop, A Cassetta, A Deacon, W N Hunter, Y P Nieh, J Raftery, N Hunter and J R Helliwell *Acta. Cryst.* **D55** 631 (1998)

[45] J R Helliwell, Y P Nieh, J Raftery, A Cassetta, J Habash, P D Carr T Ursby, M Wulff, A W Thompson, A C Niemann and A Hädener *J. Chem. Soc Farad. Trans.* **94** 2615 (1998)

[46] G V Louie, P D Brownlee, R Lambert, J B Cooper, T L Blundell, S P Wood, M J Warren, S C Woodcock and P M Jordan *Nature* **359** 33 (1992)

[47] M Helliwell, A Deacon, K J Moon, A K Powell and M J Cook *Acta. Cryst.* **B53** 231 (1997)

[48] M Helliwell (*Personal Communication*)

[49] A Yonath, J Harms, H A S Hansen, A Bashan, F Schlunzen, I Levin, I Koelln, A Tocilj, I Agmon, M Peretz, H Bartels, W S Bennett, S Krumbholz, D Janell, S Weinstein, T Auerbach, H Avila, M Piolleti, S Morlang and F Franceschi *Acta Cryst* **A54** 945 (1998)

[50] N Ban, B Freeman, P Nissen, P Penczek, R A Grausucci R Sweet F Frank, P Moore and T Steitz *Cell* **93** 1105 (1998)

[51] B T Wimberly, D E Brodersen, W M Clemons, R J Morgan Warren, A P Carter, C Vonhrhein, T Hartsch and V Ramakrishnan *Nature* **407** 327 (2000)

[52] M G Rossman *J. Synchrotron Rad.* **6** 816 (1999)

[53] E H Snell, A Cassetta, J R Helliwell, T J Boggon, N E Chayen, F Weckert, K Hölzer, K Schroer, E J Gordon and P F Zagalsky *Acta. Cryst* **D53** 231 (1997)

[54] N E Chayen, T S Boggon, J Raftery, J R Helliwell and P F Zalgalsky *Acta Cryst.* **D55** 266 (1999)

[55] N E Chayen, M Cianci, A Olczak, J Raftery, P J Rizkallah, P F Zagalsky and J R Helliwell *Acta Cryst.* **D56** 1064 (2000)

[56] M Cianci, P J Rizkallah, A Olczak, J Raftery, N E Chayen, P F Zagalsky and J R Helliwell (in preparation)

[57] E J Gordon, G A Leonard, S McSweeney and P F Zagalsky (in preparation)

[58] F Shu, V Ramakrishnan and B P Schoenborn *PNAS (USA)* **97** 3872

[59] N E Chayen, T J Boggon, A Cassetta, A M Deacon, T Gleichmann, J Habash, S J Harrop, J R Helliwell, Y P Nieh, M R Peterson, J Raftery, E H Snell, A Hädener, A C Niemann, D P Siddons, V Stojanoff, A W Thompson, T Ursby and M Wulff *Quart. Rev Biophys.* **29** 227 (1996)

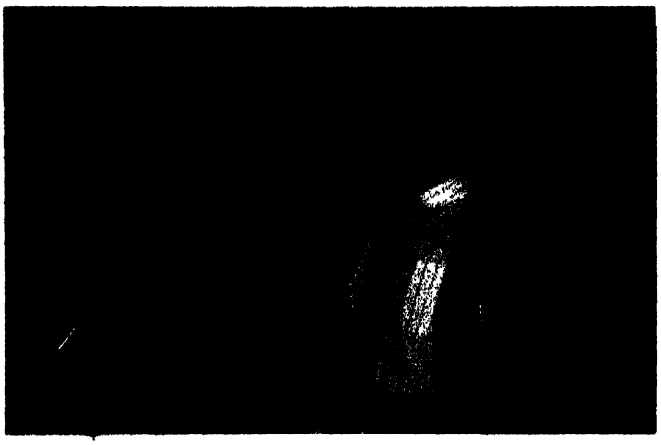
[60] J R Helliwell and M Helliwell *Chem. Comm. Feature Article* **14** 1595 (1996)

[61] E H Snell, N E Chayen, and J R Helliwell *The Biochemist.* **21** (6) 19 (1999)

[62] T J Boggon, J R Helliwell, R A Judge, A Olczak, D P Siddons, E H Snell and V Stojanoff *Acta Cryst.* **D56** 868 (2000)

[63] S E Ealick and J R Helliwell (Eds) Transactions of the American Crystallographic Association No 34. (ACA, Buffalo, NY, USA) and articles therein (1999)

The Author



John R. Helliwell, FInstP, FRSC, FInstBiol has a first class honours degree in Physics (1974, University of York), a D. Phil in Molecular Biophysics (1978, University of Oxford) and a D.Sc in Physics (1996, University of York). He is Professor of Structural Chemistry, since 1989, at the University of Manchester. He is Editor-in-Chief of Acta Crystallographica, since 1996, and which is published by the International Union of Crystallography (IUCr). His research interests span crystallographic methods especially synchrotron radiation, anomalous scattering and Laue methods, and more recently the neutron Laue method, as applied to structure determination of molecules of biological and chemical interest especially plant lectins, enzymes and, most recently, the carotenoid binding proteins responsible for the camouflage of marine crustacea. He is deeply honoured to be the first recipient of the Professor K. Banerjee Endowment Memorial Lecture Silver Medal.

Professor Helliwell proudly wearing the Endowment Medal along with Professor S P Sen Gupta (Chairman, Organizing Committee), Professor Debashis Mukherjee (Director, IACS) and Mr K C Banerji (Son-in-law of Professor K Banerjee)